

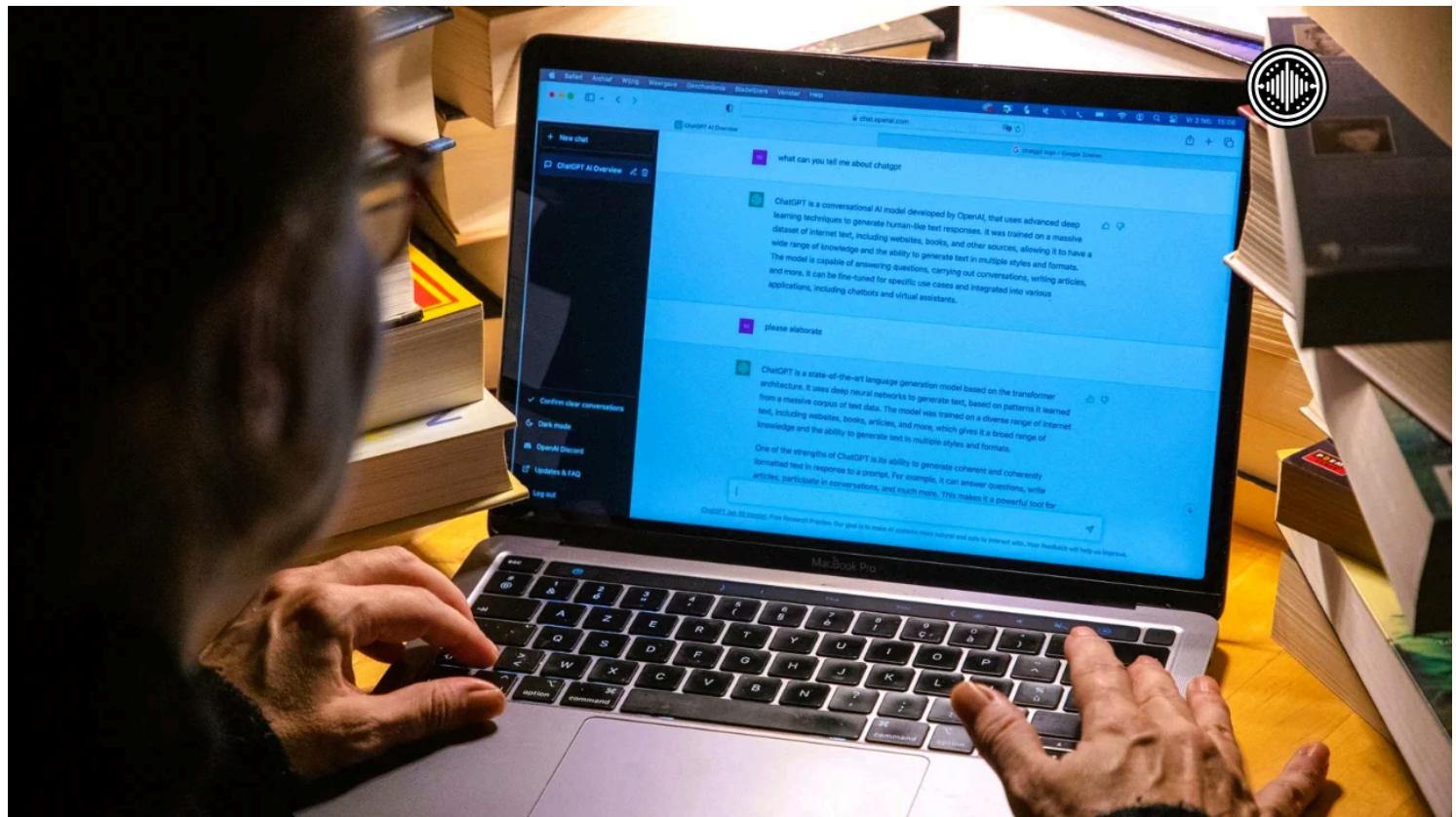
OpenVoiceNews U.S.

Transparent. Unbiased. Yours.

AI Chatbots Could Undermine Trust in Press Freedom, MIT Study Warns

July 3, 2025

— Categories: *Human Rights*



Artificial intelligence chatbots have quickly become part of daily life, answering questions, summarizing news, and even guiding decisions. However, a new study by researchers at the Massachusetts Institute of Technology Sloan School of Management suggests that these popular tools may quietly reshape how people see democracy and human rights.

The researchers examined six widely used large language models (LLMs), including ChatGPT, Gemini, and DeepSeek, to see how they assess press freedom in countries worldwide. Their findings were striking. Nearly all the models consistently painted a darker

picture of global press freedom than respected reports like the World Press Freedom Index, which is published by the non-governmental organization Reporters Without Borders.

For example, ChatGPT rated 97 percent of 180 countries negatively, implying that press freedom was weak almost everywhere. Even countries known for strong protections for journalists were often described as lacking fundamental liberties.

This misrepresentation may not be random. The study points to what the researchers call a “democratic dilemma.” Journalists and citizens can openly criticize government policies in free societies, creating more negative news coverage fed into AI training data. In contrast, authoritarian regimes often suppress critical reporting. As a result, the chatbots learn a skewed version of reality, where open societies look troubled simply because they have more visible debate and scrutiny.

Another concerning discovery was “in-group bias.” The researchers noticed that the models tended to rate their developers’ countries more favorably. For example, models created in Western nations showed a softer tone about those regions while being harsher on other countries. This built-in favoritism could have far-reaching effects, especially as AI becomes more central to news consumption and even policy-making.

“These systems have millions of users worldwide,” said Isabella Loaiza, a postdoctoral researcher at MIT Sloan and one of the study’s authors. “Their misrepresentations can distort the public’s understanding of civic rights and journalists’ freedoms.”

The implications are profound. If AI chatbots become trusted sources of information, their biases could quietly shape public opinion, undermining confidence in democratic institutions and downplaying the severity of press restrictions in authoritarian states.

Some models, like DeepSeek, have already faced criticism for censoring politically sensitive topics, especially in countries like China. Researchers warn that the risk of spreading distorted narratives grows even more urgent as AI tools are integrated into official bodies, including the United Nations.

“Access to reliable information about the institutions that uphold democracy is critical,” said co-author Roberto Rigobon. “These technologies must be aligned with democratic values, or we risk eroding the foundations of free societies.”

As AI becomes more powerful, it has never been more important to ensure that these systems reflect reality rather than amplify bias.